Influence of sampling design parameters on biomass predictions derived from airborne lidar data

M. Bouvier, S. Durrieu, Richard A Fournier, N. Saint-Geours, G. Vincent, D. Guyon, Eloi Grau, Bruno Hérault

Influence of sampling design parameters on biomass predictions derived from airborne lidar data

Marc Bouvier[1], Sylvie Durrieu[1], Richard A. Fournier[2], Nathalie Saint-Geours[1], Grégoire Vincent[3], Dominique Guyon[4], Eloi Grau[1,5], Bruno Hérault[6]

[1]*Irstea, UMR TETIS, Maison de la Télédétection, 34196 Montpellier, France*
[2]*Centre d'Applications et de Recherches en Télédétection (CARTEL), Département de géomatique appliquée, Université de Sherbrooke, Sherbrooke, Québec, Canada, J1K 2R1*
[3]*IRD, UMR AMAP, Montpellier, 34000 France*
[4]*INRA, UMR1391 ISPA, 33140 Villenave d'Ornon, France*
[5]*CNES, 31400 Toulouse*
[6]*Cirad, UMR EcoFoG, Kourou, 97300, France*

Highlights: The study provides technical guidelines to optimize sampling design parameters in the context of the implementation of biomass predictive models from airborne lidar data. Parameters under study were those easily controlled when implementing an area-based approach, i.e. lidar pulse density and both field sampling protocols and measurements. Parameters most impacting the model accuracy were identified by conducting a one-factor-at-a-time and a global sensitivity analyses. Number and size of field plot, and plot center position accuracy were identified to be the parameters most impacting model accuracy.

Keywords: **Airborne lidar, Area-based approach, Aboveground Biomass, Model accuracy, Sensitivity analysis, Monte Carlo analysis.**

Introduction

Lidar (Light Detection And Ranging) is used for many forest applications, and, in particular, to support forest inventory. Airborne laser scanning (ALS) coupled with field measurements is an effective approach that can be used to develop predictive models for assessing forest inventory attributes over large areas at a much lower cost than with traditional inventory practices. ALS is now used operationally to enhance existing inventories. With the increased use of ALS in forest applications, good survey design is increasingly important to enhance information content while maximizing cost-effectiveness. However, previous lidar studies have reported considerable variability in aboveground biomass (AGB) prediction accuracy [1]. Numerous parameters may affect the ability to reliably predict forest parameters from ALS data. Prediction accuracy on AGB primarily depends on three groups of parameters: (1) stand complexity, (2) lidar sensor and flight parameters, and (3) field protocols and measurements. These parameters affect prediction quality and consistency. Stand complexity is inherent to the sites under study and cannot be modified; one must cope with it and try to use models that have proven their effectiveness in complex environments. Nevertheless, the two other groups of parameters can be studied and carefully defined in order to maximize the chances of meeting accuracy requirements.

Deciding which lidar sensor and flight parameters are the most suitable when planning and designing a lidar survey involves a trade-off between acquisition cost and result accuracy. Pulse density (in pulses/m^2) is a key parameter in area-based approaches and a major cost driver dictated by ALS system setting and flight configuration [2]. Regarding the effects of pulse density, several studies found only small differences in results when comparing stand attributes predictions from data acquired at different pulse densities varying from 0.8 to 3.2 pulses/m² [3]. However, relevant lidar metrics selected to build predictive models were found to differ significantly with pulse density. It is therefore important to investigate how, and to what extent, pulse density affects stand attribute prediction accuracy.

Field protocols and measurements involve other parameters affecting stand attribute prediction accuracy that need to be optimized as field surveys are time consuming and costly. Field protocol design requires setting many rules regarding the choice of the number and the size of field plots, their spatial distribution, the threshold of the diameter at breast height (DBH, trunk diameter measured at 1.3 m above ground) defining trees to be inventoried, to name but a few. The cost associated with fieldwork is highly dependent on these choices. Gobakken and Næsset [4] examined both the number and size of field plots. They concluded that the optimal configuration is a tradeoff depending on inventory costs and forest structure. Plot size was also shown to influence predictions of AGB as larger plots reduce edge effects [5]. Moreover, higher prediction accuracy has been observed with an increase in plot size [1]. GPS position errors were also found to impact significantly prediction accuracy in Norway forests with a decrease in volume prediction accuracy of 15.8% for position errors up to 5 m [6]. Frazer et al. [5] investigated how plot size and co-registration errors interact to influence

AGB prediction. They found that an increase in circular plot radius from 10 to 25 m reduced the impact of co-registration error and improved AGB prediction accuracy by 13.3%. Thus, volume and AGB predictions from ALS data are highly dependent on the way field surveys were conducted as well as on the way lidar data were acquired. Both these groups of parameters can be, to some extent, set in order to optimize assessment of forest attributes from ALS data.

There is a need for more comprehensive approaches that can quantify the specific impacts of different lidar acquisition parameters, field protocols and measurements on the accuracy of the resulting model. Among the numerous parameters influencing AGB predictive model accuracy, only a few of them can be easily controlled, i.e. lidar pulse density and both field sampling protocols and measurements. Therefore, the goal of this study was to provide technical guidelines to optimize lidar pulse density and field survey protocols in order to implement predictive models of AGB from ALS data. To that aim parameters most impacting the model accuracy were identified by conducting a one-factor-at-a-time and a global sensitivity analyses.

Material and methods

Two study sites were selected: a pine plantation and a tropical forest. The pine site is located in the Landes region in southwestern France (44.40° N, 0.50° W). This forest area is dominated by mono-specific stands of Maritime pine (*Pinus pinaster* Aiton) in even aged plantations. The tropical site is a lowland rainforest located in coastal French Guiana (5.25° N, 52.93° W). Several Lidar data sets are available at several point densities on both sites (0.5 to 4.0 pulses/m² on the first site and 3.5 to 10.0 pulses/m² on the second site).

We investigated the influence of sampling design parameters on AGB prediction accuracy from ALS data using an ABA. Regression models for AGB predictions were developed according to a single method, recently developed by Bouvier et al. [7], so as to focus on the parameters explaining AGB prediction accuracy variability rather than model selection. As impacts are generally inter-dependant and thus difficult to compare, the relative importance of parameters is difficult to assess. The influence of these key parameters was assessed individually by carrying out sensitivity analysis in order to define which parameter has the most influence on AGB prediction accuracy. Thus, changing parameters one at a time (OAT approach), we explored the impact of: pulse density, the number of field plots, field plot size, minimum DBH threshold (DBH_{min}), DBH and H measurement accuracies and field plot position accuracy. As some uncertainties can accumulate and propagate, the influence of field parameters was also investigated in a Global sensitivity analysis (GSA). GSA aims to study how the uncertainty of a model output, i.e. forest attributes, can be apportioned to different sources of uncertainty in its inputs, i.e. survey parameters. It allows for a ranking of the input variables according to their contribution to the output variability. GSA thus helps to identify the key parameters which deserve specific attention. In the first study site GSA was used to study impact of four parameters - i.e., DBH and H measurement accuracies, field plot position accuracy, and allometric equations used- and to compare the relative impacts for four field protocols (two DBH_{min} and two field plot radiuses).

Results

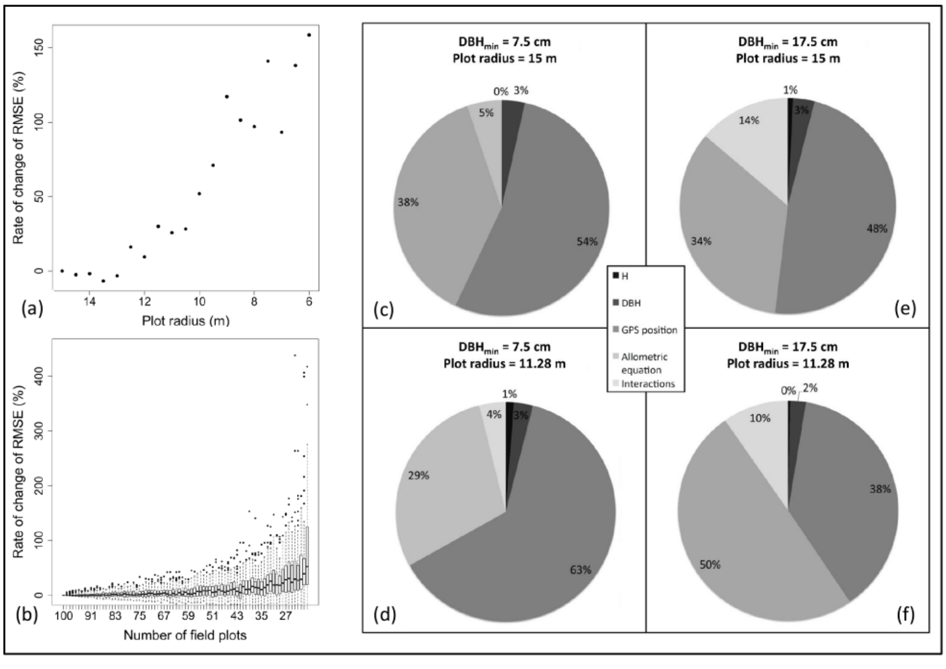

Figure. Results on pine plantation: (a) Rate of change of RMSE, expressed as a percentage of the RMSE obtained with the maximum plot radius for AGB models calibrated and validated using different field plot radius *from 15 to 6 m. (b) Rate of change of RMSE using different field plot radius from 100 to 20 m. Subsets were randomly selected from among all the field plots and selection steps were repeated 100 times. (c) - (f) Part of variance explained by H, DBH and field plot position measurement errors, and allometric equations used to predict AGB. Four protocols have been investigated: (c) 7.5 cm minimum DBH threshold (DBH_{min}) and 15 m field plot radius; (d) 7.5 cm DBH_{min} and 11.28 m field plot radius; (e) 17.5 cm DBH_{min} and 15 m field plot radius; and (f) 17.5 cm DBH_{min} and 11.28 m field plot radius.*

A one-factor-at-a-time (figure (a)-(b)) and a global sensitivity analysis (figure (c)-(f)) were applied to identify the parameters most impacting model accuracy. Variability in plot size and number of field plots were observed to significantly impact model performance (figure (a)-(b)). Preliminary results on the tropical forest site confirmed the high impact of plot size and plot number on AGB estimation accuracy. The part of variance attributable to each parameter, and linked to interactions between parameters, varied depending on which of the four inventory protocols was concerned (figure (c)-(f)). Whatever the inventory protocol, the part of variance explained by plot center positioning accuracy (38% - 63%) or the allometric equation used (29% - 50%) were significantly higher than the part of variance explained by the DBH (2% - 3%) and H (0% - 1%) parameters. The part of variance attributable to the interaction between parameters was found to vary between 10 and 14%.

Conclusion

Some recommendations can be drawn from our results for those interested in estimating AGB from a lidar-based ABA in a pine plantation. First, regarding ALS data acquisition, cost savings can be made by reducing pulse density to 0.5 pulse/m² without any major impact on model quality. Even if DBH and H measurement accuracies were shown to contribute to a lesser extent to the prediction error, field measurement costs will still remain high to ensure a good quality model. This is due to requirements regarding field plot number, position accuracy and plot size. Therefore, and despite the relative simplicity of the environment, a minimum of 40 plots is recommended. It is also recommended to inventory field plots at least 13 m in radius when plots contain trees with a DBH equal or greater to 17.5 cm in a temperate forest plantation. In addition, plot positioning efforts should be performed with great care as this parameter has a great impact even on simple and regular stands for which a position accuracy of center plots below 5 m is highly recommended. Different recommendations may be warranted when working in more complex multispecies, multistrata dense tropical forests. The preliminary results will be compared to those obtained in the tropical site. Differences and similarities between the two forest types will enable us to evaluate influence and sensitivity of design parameters with regard to the forest complexity.

References

[1] S. G. Zolkos, S. J. Goetz, and R. Dubayah, « A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing », *Remote Sens. Environ.*, vol. 128, n° 0, p. 289‑298, janv. 2013.

[2] J. Hyyppä, H. Hyyppä, D. Leckie, F. Gougeon, X. Yu, and M. Maltamo, « Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests », *Int. J. Remote Sens.*, vol. 29, n° 5, p. 1339‑1366, 2008.

[3] E. Næsset, « Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data », *Remote Sens. Environ.*, vol. 113, n° 1, p. 148‑159, 2009.

[4] T. Gobakken and E. Næsset, « Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data », *Can. J. For. Res.*, vol. 38, n° 5, p. 1095‑1109, 2008.

[5] G. W. Frazer, S. Magnussen, M. A. Wulder, and K. O. Niemann, « Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass », *Remote Sens. Environ.*, vol. 115, n° 2, p. 636‑649, 2011.

[6] T. Gobakken and E. Næsset, « Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data », *Can. J. For. Res.*, vol. 39, n° 5, p. 1036‑1052, 2009.

[7] M. Bouvier, S. Durrieu, R. A. Fournier, and J.-P. Renaud, « Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data », *Remote Sens. Environ.*, vol. 156, n° 0, p. 322‑334, janv. 2015.